

# Multiscale Analysis of Spreading in a Large Communication Network

Mikko Kivelä<sup>1,‡</sup>, Raj Kumar Pan<sup>1</sup>, Kimmo Kaski<sup>1</sup>,  
János Kertész<sup>1,2</sup>, Jari Saramäki<sup>1</sup> and Márton Karsai<sup>1</sup>

<sup>1</sup> BECS, School of Science and Technology, Aalto University, P.O. Box 12200, FI-00076

<sup>2</sup> Institute of Physics and BME-HAS Cond. Mat. Group, BME, Budapest, Budafoki út 8., H-1111

**Abstract.** In temporal networks, both the topology of the underlying network and the timings of interaction events can be crucial in determining how some dynamic process mediated by the network unfolds. We have explored the limiting case of the speed of spreading in the SI model, set up such that an event between an infectious and susceptible individual always transmits the infection. The speed of this process sets an upper bound for the speed of any dynamic process that is mediated through the interaction events of the network. With the help of temporal networks derived from large-scale time-stamped data on mobile phone calls, we extend earlier results that point out the slowing-down effects of burstiness and temporal inhomogeneities. In such networks, links are not permanently active, but dynamic processes are mediated by recurrent events taking place on the links at specific points in time. We perform a multi-scale analysis and pinpoint the importance of the timings of event sequences on individual links, their correlations with neighboring sequences, and the temporal pathways taken by the network-scale spreading process. This is achieved by studying empirically and analytically different characteristic relay times of links, relevant to the respective scales, and a set of temporal reference models that allow for removing selected time-domain correlations one by one.

*Keywords:* Network dynamics, Communication, supply and information networks, Socio-economic networks, Epidemic modelling

‡ Corresponding author: mikko.kivela@aalto.fi

## 1. Introduction

Dynamics on complex networks often take place in a non-continuous manner where interactions are not permanent. Temporal networks [1] constitute the adequate framework for such a situation, where a link between two nodes is present only for the period of the interaction. When aggregated over time, such systems can be represented with static weighted networks [2], where the link weights measure averaged link activity. Some of the properties of dynamic processes taking place on networks depend strongly on static network characteristics, while a detailed analysis of the temporal aspect could lead to further important insights.

Spreading is one of the most important processes in complex systems. It is relevant for a number of fields and applications ranging from epidemiology of biological viruses to the dynamics of social processes, such as opinion dynamics or information transmission [3]. While certain static characteristics of complex networks work to enhance spreading, such as the small world property, it has been shown that the temporal characteristics of links may slow it down [4, 5, 6]. These results indicate that dynamic processes cannot necessarily take advantage of topologically shortest paths [7]. In addition, also static topological characteristics such as prominent community structure have been shown to give rise to considerable decelerating effects on spreading speed [8, 9, 5], while a fat-tailed degree distribution has been shown to be an accelerating property [10]. Furthermore, in weighted networks, the relationship between weights and topology provides an additional source of possible influence on the spreading dynamics. Especially for social networks it is known that links within communities are strong, while links between them are weaker [11] – such Granovetterian structure enhances the trapping effect by the communities, leading to further slowing down of spreading [11, 5].

Spreading dynamics is typically studied using one of the standard epidemic models, where nodes are usually assigned to one of the three states – (S)usceptible, (I)nfectious, and (R)ecovered [12, 13, 14, 15]. The states that constitute the model are chosen depending on the problem at hand (e.g., SI, SIR, SIS). The models are then characterized by the rate of transmission between infectious and susceptible individuals upon contact, and the rate of recovery of infectious individuals. For static networks, transmission is usually assumed to take place with a probability uniform in time between infectious and susceptible network neighbors, *i.e.*, it is assumed that the timings of contacts along links are determined by a Poisson process. However, in reality the timings of such contact sequences are heterogeneous and display several kinds of temporal correlations. This is the case, e.g., in sexual contact networks [16, 17, 18] and human communication networks [19, 4, 6, 20, 5]. For both cases, it has been shown that the spreading dynamics are considerably affected by the temporal features of the contact sequences.

In this article, we employ a deterministic version of the SI epidemic model, where susceptible nodes always become infected if they are in contact with infectious nodes through a contact event, *i.e.* a phone call. We restrict ourselves to a case where only one node is initially set infected at a random point in time. Although simple, this model is useful because it gives an upper limit to the speed at which any dynamical process where nodes affect their neighbors through contacts can evolve. This includes all the other spreading processes. The setting is the same as in our previous study [5], where it was shown using temporal reference models that in addition to static structural features, heterogeneous contact sequences slow down

spreading on empirical temporal networks of mobile telephone calls and emails. We extend these results by organizing the reference model framework such that we can observe the hierarchical relationship between them. This allows us to quantify the significance of both topological and temporal correlations to the dynamic spreading speed, and also to discover their relative importance.

We perform a multiscale analysis of the effects of temporal inhomogeneities on spreading dynamics, beginning with the role of individual links and moving then on to larger scales. We zoom in into the network and focus on the effect of temporal inhomogeneities in the transmission speed of infection through single links in isolation, captured with the relay time of the link. Assuming random arrival of infection at one of the connected nodes enables to analytically calculate the effect of the inter-event time distribution [4]. It turns out that the expected transmission speed depends on the second moment of that distribution, or equivalently, on the burstiness of the sequence of events taking place on the link. However, such a simplified picture ignores correlations between such event sequences. These are taken into account in two steps. At the next scale, we consider relay times that take nearest-neighbor correlations into account with an approach similar to that discussed by Miritello *et al.* [6]. We build a model which shows that this type of correlations are likely to speed up the local spreading process. Then, we make an attempt to take into account network-wide correlations of the event sequences and measure the actual speed of transmission of infection through the links during the global scale SI process. Our investigations show that most of the temporal inhomogeneities in the slowing down of spreading can be attributed to the single link properties, and although the static link weight dominates the time the infection waits before crossing the link, there is a high variation between waiting times of links with the same weight. The inclusion of the neighbors and the effect of the whole network both lower the relay time estimates.

We begin by introducing the data set and the reference models for the contact sequence that we are going to use in all of the following sections. In Section III, we study the spreading speed at the scale of the entire network. We explore the effects of different inhomogeneities, both dynamical and topological, and show that the bursts in the event sequences of links significantly slow down the spreading – their effect is stronger than those of the weight-topology correlations or higher-order topological correlations such as communities. Finally, in Section IV we focus on the scale of links, defining the quantities that measure the spreading speed through individual links in isolation and in relationship to activation sequences by the other links.

## 2. Data and methods

### 2.1. Data

The data used in the following is based on the mobile phone call (MPC) records of seven million private customers of a single mobile network operator, covering  $\sim 20\%$  of the population of a European country. The data contain a sequence of  $\sim 444$  million time-stamped voice call events over a period of 120 days. In order to consider real social relations, we only consider communication between pairs of subscribers who have at least one reciprocated pair of calls between them. The temporal network, *i.e.* the call event sequence used in the study is constructed as follows. First, we aggregate the events over the entire period to form a weighted network, where the nodes represent subscribers and a link between two nodes indicates at least one call

both ways. For static reference models, we also compute link weights as the total number of calls made during the observation period. We then extract the largest connected component (LCC) of the aggregated network; the LCC has  $N = 4.6 \times 10^6$  nodes and  $L = 9 \times 10^6$  links. The temporal network employed in the study is then constructed as the time-stamped sequence of calls between all nodes in the LCC, and contains 306 million call records. The static LCC has broad degree distribution and shows small-world properties. Its average degree  $\bar{k} = 3.96$ , and the average shortest path length between nodes  $\langle \ell \rangle = 12.31$ . By geographic sampling of the network, we observe that  $\langle \ell \rangle$  varies logarithmically with the system size. Since the observation period is finite, we apply periodic temporal boundary conditions by repeating the call sequence once its last event is reached.

## 2.2. Reference models

For static networks, a common way to assess the significance of chosen topological features is to compare their abundance or characteristics against some reference model where the network is randomized. This approach has also been applied in assessing the importance of such features for dynamical processes. The most widely applied reference model is the configuration model [21], where the links of the original network are rewired pairwise randomly. This reference model preserves the original degree sequence but yields networks that are as random as the degree sequence allows. Then, one can assess the significance of topological characteristics of the empirical graph *e.g.* by measuring the extent to which the dynamics of some processes differ when they take place on the original networks or the reference ensemble.

For temporal graphs, a similar approach can be applied: in this case, the original event sequences are randomized or randomly reshuffled to remove time-domain structure and correlations [9, 5]. Thus a reference model for an empirical event sequence is a maximally random ensemble of event sequences, for which some predefined set of properties are the same as for the empirical sequence. There are various kinds of temporal correlations, and thus no single, general-purpose reference model (a “temporal configuration model”) can be designed. Rather, by applying appropriate reference models, one may switch off selected types of correlations in order to understand their contribution to the observed properties. The reference models used in this paper are listed in Table 1 and a short explanation is given below. For full details, see Appendix A.

*The equal-weight link-sequence shuffled model.* Whole event sequences with time stamps are randomly exchanged between links that have the same weights, *i.e.* numbers of events. Timing correlations between adjacent links are destroyed. While temporal characteristics of link event sequences are retained, any correlations between them and the topology are lost. All other temporal and structural correlations are retained.








*Link-sequence shuffled model.* As above, but sequences are exchanged between links of any weight. Thus, weight-topology correlations are additionally destroyed.

*Time-shuffled model.* The time stamps of the whole event sequence are randomly reshuffled. Thus all temporal correlations with the exception of network-level frequency envelope (for calls, the daily pattern [22]) are destroyed, while all topological features are retained.

*Uniformly random times.* Time stamps of all events are drawn uniformly from the observation period; all temporal correlations are destroyed.

*Configuration model.* Topological correlations are destroyed with the configuration model that retains the degree sequence and keeps the network connected. Original event sequences of links are then randomly redistributed. Only single-link temporal characteristics remain.

*Random network.* A connected Erdős-Rényi network is generated with the original number of nodes and links. Event sequences are distributed as above.

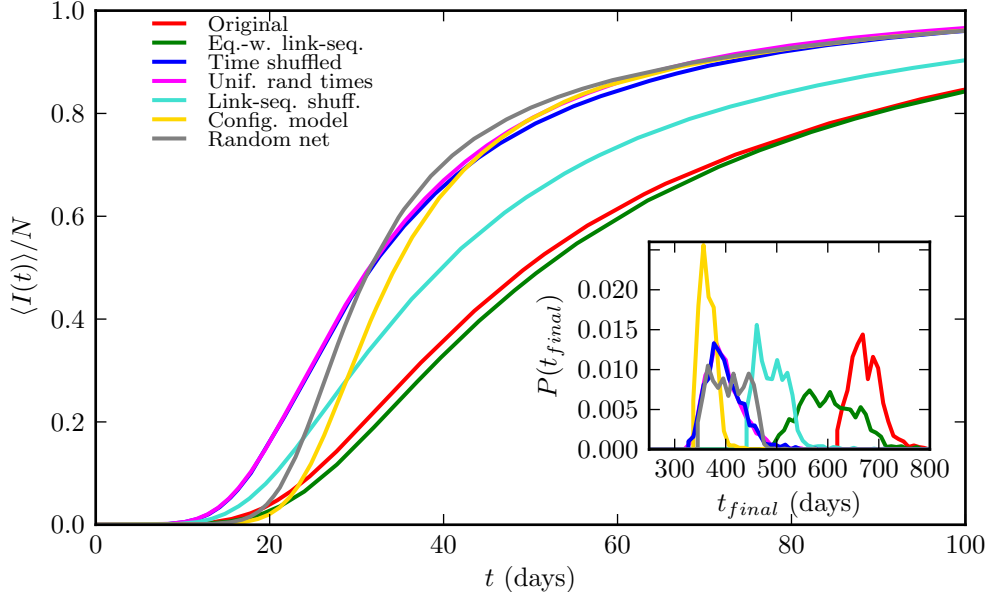
Shuffling method	Static			Temporal			Color
	DD	C	W	D	B	E	
Original	✓	✓	✓	✓	✓	✓	
Equal-weight link-sequence shuffled	✓	✓	✓	✓	✓	×	
Time shuffled	✓	✓	✓	✓	×	×	
Uniformly random times	✓	✓	✓	×	×	×	
Link-sequence shuffled	✓	✓	×	✓	✓	×	
Configuration model	✓	×	×	✓	✓	×	
Random network	×	×	×	✓	✓	×	

**Table 1.** Properties that are retained in different random reference models. Descriptions of the abbreviations are given below. The properties which can be measured from an event sequence are in parenthesis. DD - degree distribution (degree sequence), C - static configurations (adjacency matrix), W - weight-topology correlations (weight matrix), D - global time sequence (sorted list of all event times), B - event sequences in links (sorted list of even sequences on each link), E - everything (the whole event sequence). The number of links and the link weight distribution of the empirical event sequence are retained in all reference models. Further, we always require that reference networks are connected. The colors defined here for each reference model are used in the rest of the article.

### 3. Network-level spreading dynamics

We begin by simulating the SI spreading process on the empirical call sequence data and with a set of six reference models. We set one random node's state to infectious at a random point in time  $t = t_0$ , and let the information spread through the call events, measuring the time  $t_{\text{inf},i}$  from  $t_0$  at which each of the nodes gets infected. This process is repeated for  $10^3$  initial conditions. In Figure 1, we plot the average fraction of infected nodes  $\langle I(t) \rangle / N = \langle I(t; i_0, t_0) \rangle / N$  as a function of the time  $t$  from  $t_0$ , where the average is taken over the initial conditions  $i_0$  and initial infection times  $t_0$ . This quantity can also be interpreted as the probability of nodes becoming infected at times shorter than  $t$ ,  $\langle |\{i | t_{\text{inf},i} \leq t\}| / N \rangle$ . The inset of Fig. 1 displays the distributions of the times to full prevalence, *i.e.* all nodes becoming infected,  $t_{\text{final}} = \max_i t_{\text{inf},i}$ . As the spreading process itself is deterministic, the variance of these distributions originates from the random initial conditions.

First, we focus on the three temporal reference models that remove temporal correlations while leaving the static network intact (*equal-weight link-sequence shuffled*, *time-shuffled*, and *uniformly random times*). Fig. 1 shows that the spreading dynamics for those event sequences that contain bursts (the original and equal-weight link-sequence shuffled sequences) are slower than those for the reference models from which burstiness has been removed (sequences from the time-shuffled and uniformly random times models). This clearly indicates that burstiness of the event sequences



**Figure 1.** The average fraction of infected nodes for SI spreading dynamics on the MPC data and various reference models. (Inset) For each model, the distribution of the final infection time  $t_{final}$  is shown. All simulations were performed over  $10^3$  initial conditions.

slow down the spreading process significantly. Further, the dynamics for the original sequence and equal-weight link-sequence shuffled model closely resemble each other, as do the dynamics for the time-shuffled and uniformly random times models. The former similarity means that cross-link correlations have only a small influence on the speed of spreading for most of the duration of the process. The latter resemblance indicates that the daily pattern has also only minor effect on the spreading speed. However, when looking at the distributions of the full prevalence times shown in the inset of Fig. 1,  $t_{final}$ , it is seen that for long times, cross-link correlations speed up the process somewhat. The full prevalence times for the uniformly random times and the time-shuffled reference models are almost indistinguishable.

Next, we turn to those reference models that modify structural features of the static aggregated network. We find that when the network topology is retained but weight-topology correlations are removed with the link-sequence shuffled reference model, the spreading significantly speeds up compared to the original. This is because the reference model removes the known Granovetterian weight-topology correlations where weak links connecting dense communities of nodes act as bottlenecks [11]. In addition, if topological correlations such as the community structure and geographical embedding are removed with the configuration model, the dynamics of spreading become even faster with the exception of the early stage of the process. If we further remove the degree heterogeneity which is retained in the configuration model by constructing an Erdős-Renyi (ER) random networks (the random network model), there is again a small increase when compared to the configuration model. However, the configuration model is more efficient in spreading the infection than E-R networks

once the critical mass of infected nodes is reached, and when considering only the full prevalence times, the configuration model is the fastest lattice for spreading.

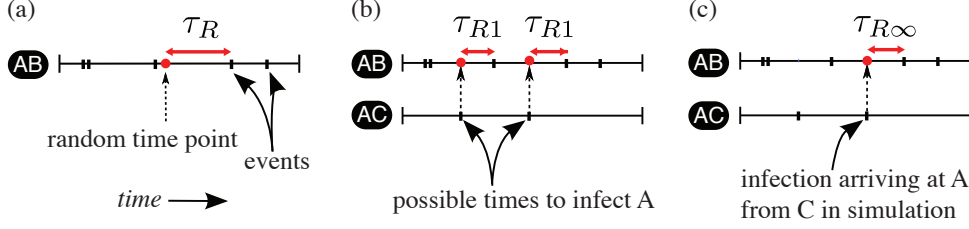
Finally, we cross-compare the relative importance of the structural and temporal correlations in the call sequence on the spreading speed. As seen in Fig. 1, the spreading dynamics for the time-shuffled model where weight-topology correlations are retained but the bursts that are destroyed are faster compared to the link-sequence shuffled model, where bursts are retained but weight-topology correlations are destroyed. This means that the burstiness of call sequences on individual links plays a more important role than weight-topology correlations in slowing down the spreading dynamics. Furthermore, the results for the configuration model and the uniform times model are roughly comparable; the first one removes all structural correlations while retaining all temporal correlations and the second one removes all temporal correlations while retaining all structural correlations. This observation can be considered surprising since the structural features of static networks are typically considered to be the most important factor for any dynamical process, while temporal correlations are assumed to have minor role. However, the above points out that the temporal correlations may be of equal importance.

In addition, we have thoroughly investigated the robustness of the above results by studying the role of the periodic boundary conditions, the initial conditions, the system size, and the error of the mean (see Appendix B and Appendix C); no significant changes to the above results were observed.

#### 4. Spreading speed on single links

In Section 3, we have shown that the temporal inhomogeneities of our empirical sequence slow down spreading on the network and that considerable amount of the slowing down can be attributed to the burstiness of event sequences of links. In order to better understand the effects of temporal inhomogeneities, we will next focus on the event sequences corresponding to each of the links, and study the effects of temporal inhomogeneities on the spreading speed at the link level. In order to do this, we investigate the statistics of the *relay times* of links. The relay time of an edge denotes the time it takes for a newly infected node to spread the infection further via the next event that the link participates in.

In the first approximation, we will neglect the effects of correlated event sequences on adjacent links and assume that nodes become infected at uniformly random times. This approximation is captured by the *random* relay time  $\tau_R$ , the waiting time between a uniformly random point in time and the next event on a link. Going beyond this approximation, nearest-neighbor correlations can be taken into account via the *triggered* relay time  $\tau_{R1}$ , where nodes may only get infected by the events they participate in. The triggered relay time is thus defined as the time between a uniformly randomly picked event involving one of the two nodes of the focal link and the next event on that link. Finally, one can consider the effects of network-wide correlations and temporal paths of the network, and measure from empirical data the actual times between the infection arriving at a node and crossing a link in simulations, leading to *actualized* relay times  $\tau_{R\infty}$ . These three concepts of relay times are schematically illustrated in Fig. 2.



**Figure 2.** Schematic illustrations of the three relay times. The horizontal line represents a link, and the small vertical lines displays the events between the two nodes within the observation period,  $t \in [0, T]$ . (a) The random relay time  $\tau_R$  is the time difference between a randomly selected point in time and the time of the next event. (b) The triggered relay time  $\tau_{R1}$  for the link  $AB$  is the time difference between the time of an event on any other link of the two nodes ( $A$  or  $B$ ) and the time of the next event on the edge  $AB$ . (c) The actualized relay time  $\tau_{R\infty}$  is calculated on the basis of simulations and measures the time between the node becoming infected and further transmitting the infection.

#### 4.1. Random relay time $\tau_R$

**4.1.1. Theory** We begin our analysis of relay times by neglecting all correlations between the event sequences of adjacent links, and study the dependence of the relay times on the inter-event time distributions of the links. We consider the case where one of the nodes,  $A$  or  $B$ , of the link  $AB$  becomes infected at a uniformly random point in time. The random relay time  $\tau_R$  is then defined as the random variable representing the time between this random time of infection and the time of the next event between  $A$  and  $B$ . In this approximation, the relay times  $\tau_R$  of a link are determined purely by the inter-event time distribution of events on that link,  $P(\tau)$ , where the inter-event time  $\tau$  is defined as the time difference between two consecutive events. On the basis of the inter-event time distribution, the probability distribution of the relay times can be written as  $P(\tau_R) = \frac{1}{\tau} \int_{\tau_R}^{\infty} P(\tau) d\tau$  [4]. Further, assuming that the tail of the distribution is not too broad ( $\lim_{x \rightarrow \infty} x^2 P(\tau > x) = 0$ ), the average relay time  $\overline{\tau_R}$  can be calculated analytically for any inter-event time distribution:

$$\overline{\tau_R} = \int_0^{\infty} \tau_R P(\tau_R) d\tau_R = \frac{1}{2} \frac{\overline{\tau^2}}{\overline{\tau}}. \quad (1)$$

The value of  $\overline{\tau_R}$  depends heavily on the average inter-event time  $\overline{\tau}$ , or equivalently, on the average rate of events taking place on the link. As we are primarily interested in the effect of the shape of the inter-event time distribution on  $\overline{\tau_R}$ , it is natural to use a Poissonian distribution as a reference. Thus, we define the normalized average relay time,  $\overline{\tau_R}^*$ , by dividing  $\overline{\tau_R}$  by the corresponding relay time for a Poisson process  $\overline{\tau_{R,p}}$ , for which the average inter-event time is matched to that of the given inter-event time distribution such that  $\overline{\tau_p} = \overline{\tau}$ . The normalized relay time  $\overline{\tau_R}^*$  can then be written as

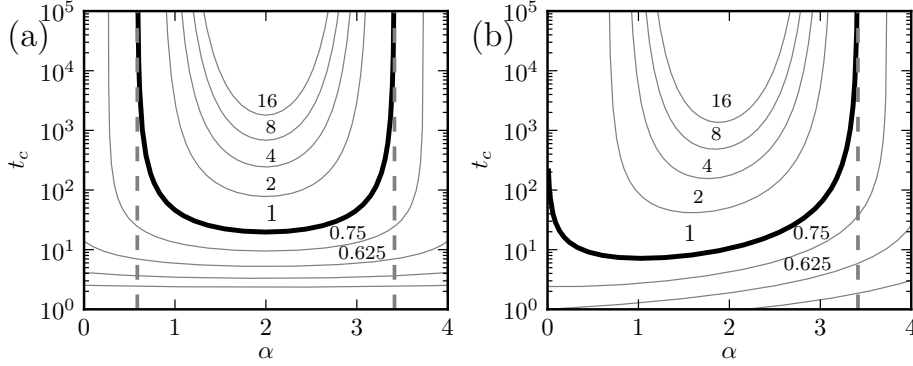
$$\overline{\tau_R}^* \equiv \frac{\overline{\tau_R}}{\overline{\tau_{R,p}}} = \frac{\overline{\tau^2}}{\overline{\tau_p^2}} = \frac{\overline{\tau^2}}{2\overline{\tau}^2}, \quad (2)$$

where  $\overline{\tau_{R,p}} = \overline{\tau_p} = \overline{\tau}$ . Thus,  $\overline{\tau_R}^*$  measures the ratio of the second moment to the square of the first moment of the inter-event time distribution. Generally, the broader



$\sim p(\tau)$	$\overline{\tau_R}$	$\overline{\tau_R}^*$	Definitions
$e^{-\tau/t_c}$	$t_c$	1	$t_c > 0$
$\tau^{-\alpha}, 1 < \tau$	$\frac{1}{2} \frac{\alpha-2}{\alpha-3}$	$\frac{(\alpha-2)^2}{2(\alpha-1)(\alpha-3)}$	$\alpha > 3$
$\tau^{-\alpha}, 1 < \tau < t_c$	$\frac{\gamma(3)}{2\gamma(2)}$	$\frac{\gamma(1)\gamma(3)}{2\gamma(2)^2}$	$\gamma(k) = \frac{1-t_c^{k-\alpha}}{\alpha-k}, \alpha \neq k$
$\tau^{-\alpha} e^{-\tau/t_c}, 1 < \tau$	$\frac{t_c}{2} \frac{\theta(3)}{\theta(2)}$	$\frac{\theta(1)\theta(3)}{2\theta(2)^2}$	$\theta(k) = \Gamma(k - \alpha, t_c^{-1})$

**Table 2.** Functional forms of the random relay times  $\overline{\tau_R}$  and the normalized random relay time  $\overline{\tau_R}^*$  for some inter-event time distributions.



**Figure 3.** Contour plots the value of  $\overline{\tau_R}^*$  when the inter-event time distribution is (a) power-law with exact cutoff and (b) power-law with exponential cutoff (See the Table 2). The black solid contour lines correspond to values of  $\overline{\tau_R}^* = 1$ , the gray  $n$ 'th line above that to  $\overline{\tau_R}^* = 2^n$  and the gray  $n$ 'th line below the black line to  $\overline{\tau_R}^* = \frac{1}{2}(1 - 2^{-n})$ . The gray dotted vertical lines are at  $\alpha = 2 \pm \sqrt{2}$ .

the distribution, the larger the second moment as compared to the square of the first moment. Because of this,  $\overline{\tau_R}^*$  can be viewed as a measure of the broadness of the inter-event time distribution. Since a broad inter-event time distribution is also a signature of a bursty time sequence,  $\overline{\tau_R}^*$  can equally well be seen as a measure of burstiness. This is further illustrated by the following:  $\overline{\tau_R}^*$  can be written as  $\overline{\tau_R}^* = [(\frac{\sigma_\tau}{\overline{\tau}})^2 + 1]/2$ , where  $\sigma_\tau$  is the standard deviation and  $\frac{\sigma_\tau}{\overline{\tau}}$  the coefficient of variation of the inter-event time distribution. The coefficient of variation  $\frac{\sigma_\tau}{\overline{\tau}}$  was also used by Goh and Barabási [23] as a basis for their measure of burstiness,  $B = (\frac{\sigma_\tau}{\overline{\tau}} - 1)/(\frac{\sigma_\tau}{\overline{\tau}} + 1)$ . Thus, the normalized average relay time can be written as  $\overline{\tau_R}^* = \frac{B^2 + 1}{(B - 1)^2}$ . Hence, the burstier an event sequence is, the longer the random relay times are.

The inter-event time distributions of empirical event sequences are often broader than would be expected from a Poisson process. Such empirical distributions have been fitted with power laws [24, 25] or power laws with an exponential cutoff [4, 26]. In Table 2 we present analytical formulas for  $\overline{\tau_R}$  and  $\overline{\tau_R}^*$  for these distributions and illustrate their dependence on the distribution parameters. For the power-law distribution,  $\overline{\tau_R}^*$  becomes infinite when the power-law exponent  $\alpha = 3$ , and decreases with increasing  $\alpha > 3$ , reaching  $\overline{\tau_R}^* = 1$  when  $\alpha = 2 + \sqrt{2}$ . Thus, for power-law inter-event time distributions in this regime, the average random relay times are longer than for the Poissonian reference case. For power-law distributions with cutoff, the cutoff value  $t_c$

also affects  $\overline{\tau_R}^*$ . The joint effect of the power-law exponent and the cutoff is illustrated with contour plots in Fig. 3. If the cutoff point  $t_c$  is small enough, the random relay times are shorter than Poissonian also for power-laws with  $\alpha \leq 2 + \sqrt{2}$ , but for large values of  $t_c$ , the power-law inter-event times dominate and the relay times become longer.

*4.1.2. Relay times in empirical event sequences* Next, we turn our attention from the general case to empirical data, where we have a finite number of events,  $n$ , in a finite observation period,  $t \in [0, T]$ . For simplicity, we consider the case with periodic temporal boundary conditions, where all events get repeated with a periodicity of  $T$ . Then, the expression of relay time distribution of Eq. (1) becomes  $P(\tau_R) = \frac{1}{T} \sum_{i=1}^n [\tau_R \leq \tau_i]$ , where  $\tau_i$  is the inter-event time between the  $i-1$ -th and  $i$ -th events. The expected relay time  $\overline{\tau_R}$  is then given by

$$\overline{\tau_R} = \frac{\sum_{i=1}^n \tau_i^2}{2T}. \quad (3)$$

For a given values of  $n$  and  $T$ , the average relay time  $\overline{\tau_R}$  is minimized when all the inter-event times are the same, that is,  $\tau_i = \tau_j, \forall i, j$ , giving  $\overline{\tau_R} = \frac{T}{2n}$ . In contrast, it is maximized when all the events happen at the same time resulting in a single non-zero inter-event time, that is,  $t_i = t_j, \forall i, j$  and  $\tau_1 = T$  giving  $\overline{\tau_R} = T/2$ . Thus,  $\overline{\tau_R}$  is maximized when all events take place in the same burst and  $\overline{\tau_R}$  is minimized when events are maximally separated in time.

As with the general case, we want to compare the  $\overline{\tau_R}$  values to some baseline. Here, the proper baseline is to distribute the times of the  $n$  events of each edge at uniformly random within the observation period, equivalently to the uniform random times reference model defined earlier. This is equivalent to having a Poisson process with a constant average inter-event time with the additional condition that there are  $n$  events in the interval [27]. Order statistics can be used to analytically derive the inter-event time distribution, which in this case is a beta distribution [28] (see Appendix D for details). The beta distribution converges to an exponential distribution as the number of events  $n$  increases, and thus the uniformly random times reference model approaches the Poisson process with increasing  $n$ . The expected value and variance of  $\overline{\tau_R}$  for the uniform times model with periodic boundary conditions are given by

$$\langle \overline{\tau_R} \rangle = \frac{T}{n+1}, \text{Var}(\overline{\tau_R}) = \frac{T^2(n-1)}{(n+1)^2(n+2)(n+3)}, \quad (4)$$

and the normalized relay time (similarly to Eq. 2) becomes

$$\overline{\tau_R}^* = \frac{\overline{\tau_R}}{\langle \overline{\tau_R} \rangle} = \frac{(n+1)}{2} \sum_{i=1}^n (\tau_i/T)^2. \quad (5)$$

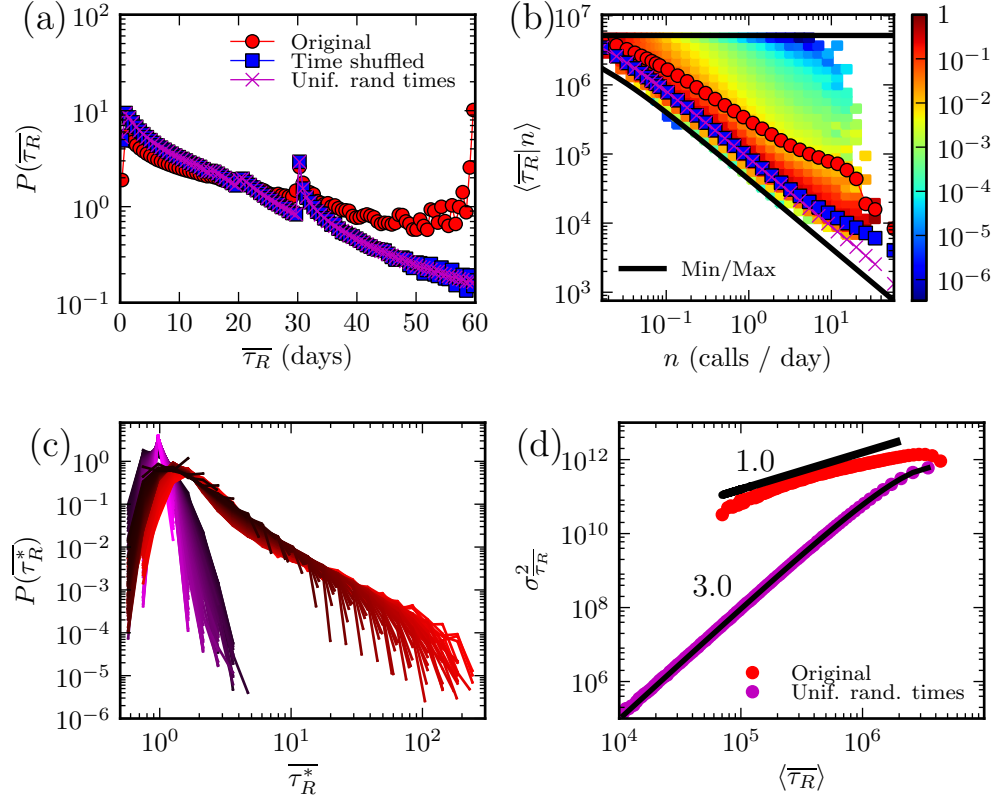
Further, since the normalization of  $\overline{\tau_R}^*$  is done with the uniformly random times reference model instead of a Poisson process, we need to include a correction factor in the relationship between burstiness and the normalized relay time:  $\overline{\tau_R}^* = \frac{n+1}{n} \frac{B^2+1}{(B-1)^2}$ .

Next we calculate  $\overline{\tau_R}$  for each edge for the empirical MPC sequence using Eq. (3) and compare the values to the temporal reference models, where the static network remains unchanged. Further, we leave out the equal-weight link-sequence shuffled model which would give equivalent results to the empirical sequence. This leaves us

only with the time-shuffled and the uniformly random times reference models. In Fig. 4 (a) we plot the probability distribution of  $P(\overline{\tau_R})$  for the original data and for the random references. It is seen that in the empirical data there are more edges with high random relay times and fewer edges with low relay times than in the reference models. This is also reflected in the value of the network-level average of  $\overline{\tau_R}$  for the original sequence, averaged over all edges: it is 23.1 days while for the reference models the average value is only 13.8 days. Thus, the high average relay times of the empirical network can at least partly explain the slowing down of the SI-process. Further, the distributions of  $\overline{\tau_R}$  for the time-shuffled and uniformly random model are indistinguishable, as expected.

The random relay times of edges depend heavily on their weight  $n$ , *i.e.* the total number of calls on each edge. Thus the shapes of the  $\tau_R$  and  $\overline{\tau_R}$  distributions are dominated by the heterogeneous weight distribution. In Figure 4 (b) we plot  $\overline{\tau_R}$  as a function of the edge weight. The difference between the original event sequence and the reference sequence grows with the edge weight, and the edges on the original sequence are on average slower than on the references independent of the weight. However, for links of the same weight, the  $\overline{\tau_R}$  values for the empirical data are broadly distributed, and the tail of this distribution is partly responsible for the longer average relay times. Note that for the time-shuffled reference model, the average relay time values are almost the same as for the uniform random times model except for very high weights, where the time-shuffled reference is slower. As other temporal correlations than the daily pattern are absent from this reference, any deviation is due to the daily pattern. It is seen that the daily pattern starts to have an effect only when the call frequency is larger than a few calls per day. More than 99% of the edges in the network have less than three calls per day, and thus for a vast majority of the edges, the daily pattern has practically no effect on the relay times, in line with our observations on spreading dynamics.

As seen in Figure 4 (b), the random relay times display a lot of variation even for edges of the same weight. We illustrate this variation in more detail in Figure 4 (c), and show the distributions of  $\overline{\tau_R^*}$  for edges with fixed weight. For high-weight edges, we use a weight range to improve statistics. Remarkably, the distributions for the original event sequence collapse together for all weights, with the exception of a small offset in the peak and differences in cutoff values ( $\max(\overline{\tau_R^*}) = (n + 1)/2$ ). Thus there is a signature distribution of the relay times that does not depend on edge weight – as seen in comparison with the uniformly random times reference model, this distribution is broad and the relay times contain a lot of variation. Thus, while many edges have relay times close to the reference model, there is a small number of highly bursty edges with very long relay times. Furthermore, the length of the tail of the  $\overline{\tau_R^*}$  distributions grows with the edge weight for the original sequence, which is in contrast to the uniformly random times sequence for which the variance is smaller for higher weights. To better quantify the difference in relay times within the group of edges with equal weight in Figure 4 (d) we show a scatter plot where each point corresponds to the variance and the average of  $\overline{\tau_R}$  for an edge group (similar to fluctuation scaling [29]). For the uniformly random times we can use Eq. (4) to calculate the relationship between the average and the variance. For large weights we get  $\langle \overline{\tau_R} \rangle \propto (\sigma_{\overline{\tau_R}}^2)^\alpha$ , with  $\alpha = 3$ . However, for the original sequence the scaling exponent is close to  $\alpha = 1.0$  and the variance is higher than for the reference model for all values of  $\langle \overline{\tau_R} \rangle$ .



**Figure 4.** (a): Distribution of the relay times  $\bar{\tau}_R$  for the original, time-shuffled and uniformly random event sequences. The peak at  $\sim 30$  days is due to the larger amount of edges with only 2 calls (a result of the mutuality condition) and periodic temporal boundary conditions. (b):  $\langle \bar{\tau}_R | n \rangle$  for the original event sequence and reference models. For the original sequence, we also show the probability distributions of  $\bar{\tau}_R$  conditional to weight (see color bar on right). The solid lines denote (theoretical) max and min values. (c): The distributions of  $\bar{\tau}_R^*$  conditional to (logarithmically binned) weights for the original sequence (red) and the uniformly random times model (magenta). Each curve is colored according to the weight associated to it such that that dark colors indicate small weights. (d): The average and the variance of  $\bar{\tau}_R$  for each of the bins. The solid line on the top indicates power law fit to the original data, and the lower black line corresponds to the analytical values for the uniformly random time sequence calculate with Eq. 4. Distributions for bins with more than 800 calls during the time interval  $T$  are not shown due to insufficient statistics.

#### 4.2. Triggered relay time $\tau_{R1}$

**4.2.1. Theory** For the random relay times  $\tau_R$  considered in the previous section, it was assumed that the event sequences on adjacent links are uncorrelated, and thus the infection may arrive at one of the nodes of the focal link at a random point in time. We will next take into account the effects of correlations between the times of events of adjacent links, and require that a node may only become infected when it interacts with one of its neighbors. Instead of choosing the time of infection of node

$A$  or  $B$  of the link  $AB$  randomly from the time interval  $[0, T]$ , we pick it randomly from the set of times where  $A$  or  $B$  participates in an interaction event with one of its neighbors  $C \neq B$  and  $C \neq A$  [6]. Then, the *triggered* relay time  $\tau_{R1}$  for the link  $AB$  is defined as the time difference between the infection arriving  $A$  (or  $B$ ) to the time when the infection is passed through the link to  $B$  (or  $A$ ), similarly to Ref. [6].

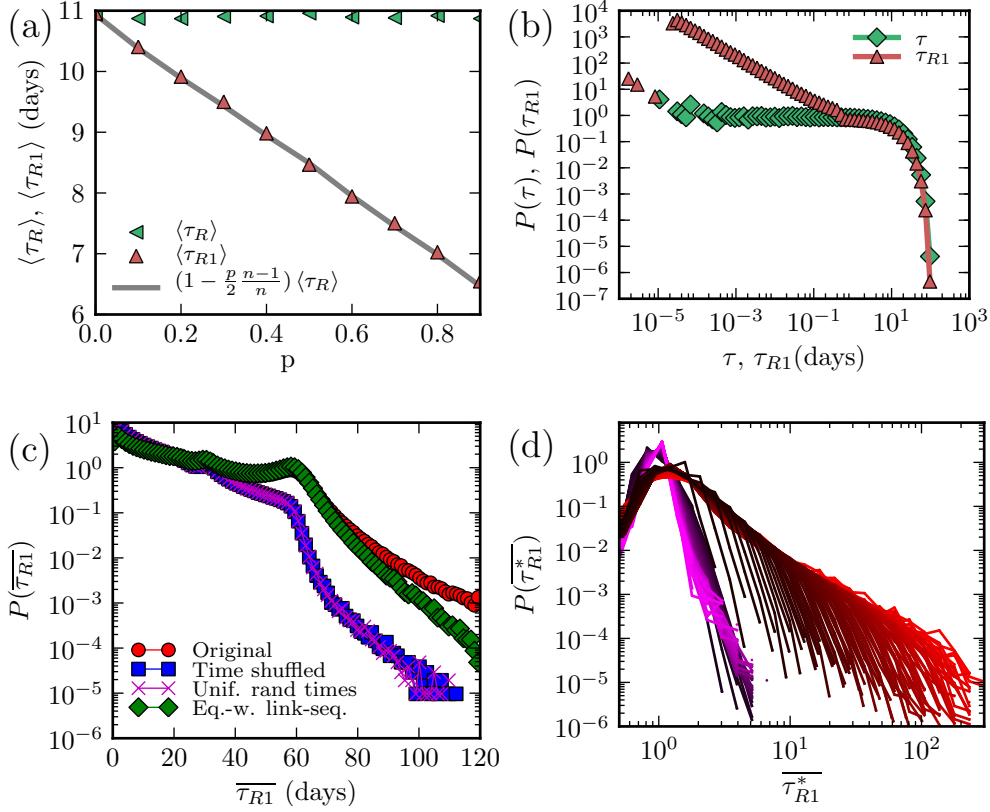
The average triggered relay time  $\overline{\tau_{R1}}$  for the link  $AB$  depends on both the time intervals between the events on the link, as well as correlations of the times of events between the link and its neighbors. If the events on the other links are independent of the events on  $AB$ , then  $\overline{\tau_{R1}}$  approaches  $\overline{\tau_R}$ . We illustrate the differences between  $\tau_{R1}$  and  $\tau_R$  with the help of a simple model that involves correlations between the sequences. We define a *triggering model* where the system consists of two links,  $AB$  and  $AC$ , such that each of the two links has  $N$  events. We assign times to the events on both links in a pairwise fashion, such that each  $AB$ - $AC$  event pair is considered to be triggered with a probability  $p$  and uncorrelated with probability  $(1 - p)$ . For uncorrelated event pairs, we assign times drawn uniformly at random from the interval  $[0, T]$ . For correlated event pairs, we first choose the time  $t_{AB}$  of the event on the  $AB$  link uniformly at random. Then we set an event on the  $AC$  link randomly at either  $t = t_{AB} - \delta t$  or  $t = t_{AB} + \delta t$ , i.e. the  $AC$  event is considered to trigger the  $AB$  event or vice versa. The time  $\delta t$  is picked at random from some triggering time distribution. Periodic temporal boundary conditions are invoked if the  $AC$  event falls outside the interval  $[0, T]$ . Thus, on both links we have on average  $pN$  events that are triggered or trigger another event, and  $(1 - p)N$  uncorrelated events.

Now we can calculate for the expected triggered relay time  $\langle \overline{\tau_{R1}} \rangle$  for the  $AB$  link by assuming that the trigger times  $\delta t$  are small compared to the random inter-event times, i.e. we always have  $\tau > \delta t$ . Note that the model is symmetric in that  $\langle \overline{\tau_{R1}} \rangle$  is same for both links,  $AB$  and  $AC$ . The fraction of uncorrelated events is  $(1 - p)$ , and for these, the expected triggered relay time is simply given by Eq. (4):  $\langle \overline{\tau_{R1}} \rangle = T / (n + 1)$ . The fraction of triggered event pairs is  $p$ ; half of the events are triggered by an  $AC$  event with a triggered relay time  $\langle \delta t \rangle$ . The rest trigger  $AC$  events, and the expected time from the triggered  $AC$  event to the next  $AB$  event is  $\langle \tau - \delta t \rangle = \frac{T}{n} - \langle \delta t \rangle$ . Thus, considering all these cases, the expected triggered relay time

$$\begin{aligned} \langle \overline{\tau_{R1}} \rangle &= (1 - p) \frac{T}{n + 1} + \frac{1}{2} p \langle \delta t \rangle + \frac{1}{2} p \left( \frac{T}{n} - \langle \delta t \rangle \right) \\ &= \left( 1 - \frac{p}{2} \frac{n - 1}{n} \right) \langle \overline{\tau_R} \rangle. \end{aligned} \quad (6)$$

Thus, the more triggered events there are, the shorter the triggered relay times are on average, as expected. Note that in the above equation  $\langle \overline{\tau_{R1}} \rangle$  does not depend on the distribution of  $\delta t$  as long as the values of  $\delta t$  are small enough compared to the inter-event times.

The behavior of the triggering model is illustrated in Fig. 5 (a), where the average random and triggered relay times from simulations of the model are shown; the latter are seen to match well with the analytical prediction of Eq. (6). For these simulations, we have used a power-law distribution of the triggering times  $\delta t$  with an exponent of 0.9 and a cutoff at  $t_c = 10^5$  seconds; this has been motivated by the shape of the low- $\Delta t$ -region of triggered relay time distributions in experimental data (see, e.g. [6]). Fig. 5 (b) shows the distribution of the inter-event times  $P(\tau)$  and the relay times  $P(\tau_{R1})$  for the model with this triggering time distribution. The distribution  $P(\tau_{R1})$  shows an initial power-law decay which is caused by the triggered events, followed by exponential decay. In contrast, the inter-event time distribution  $P(\tau)$  is exponential.



**Figure 5.** (a)-(b) Results of the triggering model with parameters  $T = 120$  days,  $N = 10$ , while  $P(\delta t)$  follows a power-law with exponent 0.9 and cutoff at  $10^5$  seconds. The simulations are performed over  $10^4$  independent realizations. (a) The distribution of  $\langle \tau_R \rangle$ , and  $\langle \tau_{R1} \rangle$  as a function of the coupling parameter  $p$ . The solid curve represents Eq. (6). (b) The distributions of  $\tau_{R1}$  and inter-event time  $\tau$  for the coupling parameter  $p = 0.4$ . Inter-event times are exponentially distributed whereas the distribution of  $\tau_{R1}$  follows a power-law with exponential tail. (c) Distributions of the average relay time in empirical data,  $P(\overline{\tau_{R1}})$ , for the original event sequence and the temporal reference models. (d) The distributions of  $\overline{\tau_{R1}}^*$  conditional to (logarithmically binned) weights for the original sequence (red, curves on the right) and the uniformly random times model (magenta, curves on the left). Each curve is colored according to the weight associated to it such that dark colors indicate small weights.

**4.2.2. Empirical results** Next we study the averages of triggered relay times  $\overline{\tau_{R1}}$  in empirical data. In Fig. 5 (c) we plot the distributions of  $\overline{\tau_{R1}}$  of the MPC network original call sequence and the temporal reference models. Similar to random relay times, the distributions of  $\overline{\tau_{R1}}$  for the time-shuffled and uniform random time model match with each other. Further, the distribution for the original event sequence has a fatter tail than these two reference models. Unlike with the case of  $\overline{\tau_R}$ , the distributions of  $\overline{\tau_{R1}}$  are not same for the original and the equal-weight link-sequence shuffled event sequences. This is because equal-weight link-sequence shuffling destroys the correlations between events on neighboring edges. In the coupling model, such

correlations caused the expected  $\overline{\tau_{R1}}$  values to decrease. This is true also for the MPC network, since the average  $\overline{\tau_{R1}}$  for the original sequence is 22 days, where it is 23 days for the equal-weight link-sequence shuffled event sequence.

In Fig. 5 (d) we plot the  $\overline{\tau_{R1}^*}$  distributions conditional to the edge weight similar to Fig. 4 (c). The main difference to  $\overline{\tau_R^*}$  distributions is that the theoretical maximum value for  $\overline{\tau_{R1}^*}$  is  $n + 1$  instead of  $(n + 1)/2$ . The values of  $\overline{\tau_{R1}^*}$  inside the interval from  $(n + 1)/2$  to  $n + 1$  arise from situations where the triggering events take place close to the beginning of a long interval between consecutive bursts. Thus, they result from a combined effect of burstiness and timings of triggering events. Such situations are not very frequent, leading to steep decays in the tails of the  $\overline{\tau_{R1}^*}$  distributions for values larger than  $(n + 1)/2$ . Thus, even for small edge weights, the tails of the  $\overline{\tau_{R1}^*}$  distributions do not follow the signature scaling observed for  $\overline{\tau_R^*}$  distributions.

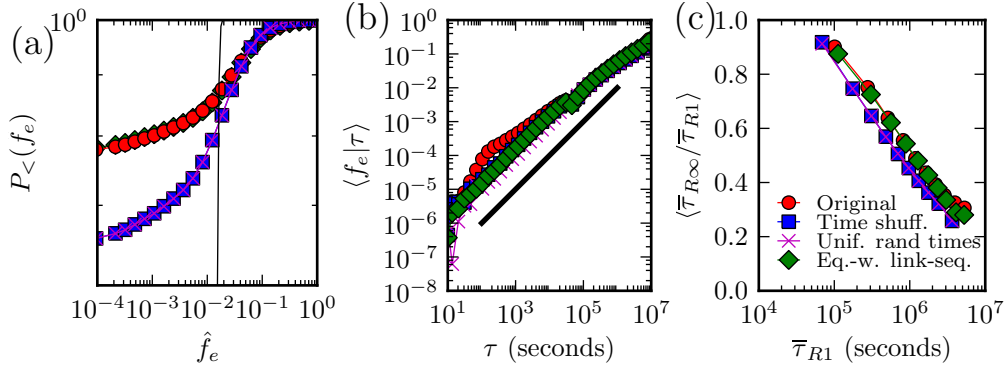
#### 4.3. Actualized relay time $\tau_{R\infty}$

**4.3.1. Motivation: the event infectivity  $f_e$**  When defining the triggered relay time  $\tau_{R1}$ , it was assumed that each event taking place on the neighboring edges transmits the infection to the focal edge with an equal probability. In this section, we study the validity of this assumption by defining a quantity called *infectivity*,  $f_e$ , for each event. The event infectivity is defined as a probability of an event spreading the infection during a spreading process which is initiated from a random node and a random point in time.

In order to calculate exact values for the event infectivities, we would need to simulate the spreading process over the entire ensemble of all possible initial conditions. Instead, we calculate an estimate for  $f_e$  by starting the infection from  $10^4$  randomly selected initial conditions and simulating each spreading process until all the nodes in the network are infected. For these processes, we calculate the number of times each event transfers the infection to a susceptible node. An estimate for the event infectivity,  $\hat{f}_e$ , is then determined by dividing these numbers by the number of initial conditions. Note that this approximation is not very accurate for small values of  $f_e$ .

In Fig. 6 (a) we plot the cumulative distributions of the estimated event infectivities  $P_{<}(\hat{f}_e)$  for the original sequence and the temporal reference models. We find that there are large differences in the number of non-infecting events (events that have  $\hat{f}_e = 0$ , *i.e.* were never observed to transmit the infection). The total fraction of events which do not participate in the spreading process is  $\sim 77\%$  for the original sequence and the equal-weight link-sequence shuffled model, while for the time-shuffled and the uniform time model it is 65%. Further, if one assumes that all the “real” event infectivity values would be equal, the sampled event infectivity estimates would be distributed with a mean  $\langle \hat{f}_e \rangle = 0.015$  and a standard deviation  $\sigma(\hat{f}_e) = 10^{-3}$  §. The observed variation in the estimated infectivities is much larger than this. Thus we can conclude that the assumption of all events being equally probable carriers of infection is not valid – there is considerable variation that arises from the timings of the events.

§ If there are  $N$  nodes,  $E$  events and the infection is started with  $L$  initial conditions, then for each event, the number of times it participates in the spreading,  $C$ , is binomially distributed with parameters  $p = \frac{N-1}{E}$  and  $n = L$ . Then, for that event  $f_e = C/L$ . Thus  $\langle \hat{f}_e \rangle = p$  and  $\sigma(\hat{f}_e) = \sqrt{\frac{p(1-p)}{L}}$ . The  $\hat{f}_e$  values for the events are not independent of each other. However, the correlations between them are weak if  $N \ll E$ , and the  $\hat{f}_e$  distribution gotten from the sampling process can be approximated by the distribution that a single event would have in this process.



**Figure 6.** (a) The cumulative distribution of estimated event infectivity,  $P_{<}(\hat{f}_e)$ , for the original sequence and the temporal reference models. The black line corresponds to the mean of the estimated infectivity distribution for a constant “real” event infectivity; the variation arising from sampling is small enough to render the standard deviation of this distribution negligibly small. (b) The average event infectivity as a function of the trailing inter-event time (time to last event). The solid black line corresponds to a linear increase. (c) The average ratio  $\langle \bar{\tau}_{R\infty} / \bar{\tau}_{R1} \rangle$  as a function of  $\bar{\tau}_{R1}$ . The edges are placed into bins according to their  $\bar{\tau}_{R1}$  values in a way that each bin contains equal number of edges. The coordinate of the bin in the x-axis is then calculated as the average  $\bar{\tau}_{R1}$  value of the edges inside the bin.

To better understand which events are most or least likely to spread the infection, we plot in Fig. 6 (b) the average estimated event infectivity as a function of the preceding inter-event time of the event, *i.e.* the time to the previous event. On average, the infectivity of an event increases with the inter-event time. This means that events participating in bursts (events with short preceding inter-event times) are less likely to participate in the infection spreading process as compared to events with long preceding inter-event times. The increase in  $\hat{f}_e$  with  $\tau$  shows almost a linear relationship for the temporal reference models; a linear relationship would be expected if the infection times were random. This is simply because if we select a random infection point in time, the probability  $f_e$  of the infection going through an event  $e$  would be proportional to the preceding inter-event time of that event  $\tau_e$ , that is  $f_e \propto \tau_e / T$ . However, the curve for the original event sequence shows a slight deviation from this linear dependency for short inter-event times.

**4.3.2. Definition of  $\tau_{R\infty}$  and empirical results** The above observations on large variations in event infectivity and correlations between preceding inter-event times and infectivities are not consistent with the assumptions behind the definition of the triggered relay time  $\tau_{R1}$ , where it was assumed that the infection can be received through each event with equal probability. Further,  $\tau_{R1}$  doesn’t take into account the possibility of both nodes of the focal edge getting infected independently, without transmission through that edge, as the spreading front expands through their neighboring nodes.

In order to account for the infectivity distributions and all other constraints during the spreading process we define the *actualized relay time*  $\tau_{R\infty}$  as the time



that the infection actually waits on one of the nodes of an edge before crossing it in a network-wide spreading process, *i.e.* the difference between the time when the infection actually arrives at one of the endpoint nodes of an edge and the infection being transmitted through that link. The values for  $\tau_{R\infty}$  are obtained numerically by running the spreading process through the network, monitoring which nodes infect which other nodes, and calculating the relay times accordingly. Note that only events transmitting the infection are considered for  $\tau_{R\infty}$ . As an example, if node  $u$  is infected and the next event between  $u$  and  $v$  occurs after node  $v$  has already become infected by some other node  $w$ , that event doesn't affect the  $\tau_{R\infty}$  value of the edge linking  $u$  and  $v$ .

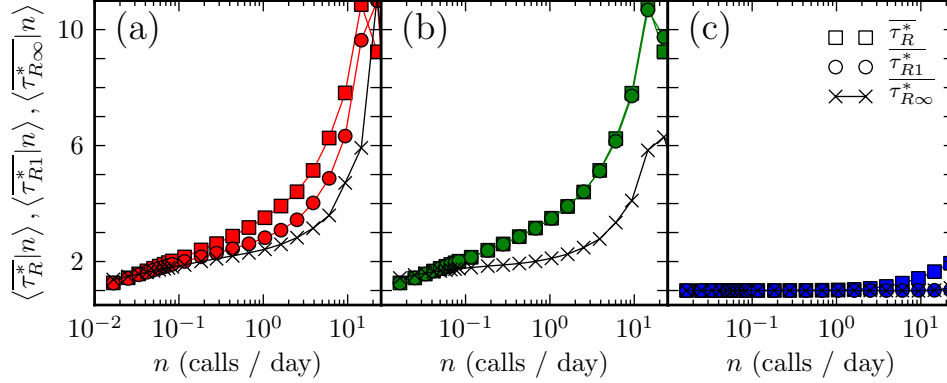
For estimating  $\tau_{R\infty}$  in empirical data, we use the same sampling scheme as previously used while estimating the infectivity values  $f_e$ . In Fig. 6 (c) we plot the ratios  $\langle \overline{\tau_{R\infty}} / \overline{\tau_{R1}} \rangle$  as functions of the respective relay times, for the original sequence and the reference models. On average, the actualized relay times  $\overline{\tau_{R\infty}}$  are shorter than the triggered relay times, and the ratio decreases with increasing  $\overline{\tau_{R1}}$ . Thus, when the network-wide spreading process is accounted for, the relay times speed up especially for edges that appear slow when measured with triggered relay times. This is because the pathways taken by the spreading process often avoid and circumvent the slowest edges. However, in some cases it is not possible to employ any faster shortcuts, *e.g.* an edge leading to a node of degree one is always used in the spreading process. Note that  $\overline{\tau_{R\infty}}$  is not defined for edges that only have zero infectivity events, *i.e.* for edges that never transmit the infection.

#### 4.4. Comparing the relay times

Having introduced the three different relay times and illustrated their differences, we now turn to comparing them in detail with regards to the empirical MPC data. For the comparison, we normalize the average relay times with the expected times for the uniformly random times reference model. The normalized random, triggered, and actualized relay times as a function of edge weight are shown in Figs. 7 (a), (b) and (c) for the original event sequence, equal-weight link-sequence shuffled reference and time-shuffled reference, respectively.

Let us first focus on the scale of individual links and the behavior of the random relay times  $\tau_R$  that do not take correlations between links into account. For the original sequence, the random relay times are always longer than for the uniformly random sequence, even for the smallest weights. Further, the difference grows with edge weight, reaching a maximum at  $n \sim 12$  calls per day where the random relay times of the original sequence are on average 11 times longer than for the reference. These effects reflect the burstiness of call sequences of edges. For the equal-weight link-sequence-shuffled reference, the random relay times are necessarily equal to the original sequence. For the time-shuffled reference, the random relay times deviate from the uniform case only for the highest-weight links. This means that the daily pattern has some effect on the highest-weight links, because it is the only difference between the two reference models; for the majority of links, it has no visible effect.

Next, we observe how the relay times change when information about events in the local neighborhood of the focal edge is taken into account with the triggered relay times  $\tau_{R1}$ . For the original event sequence, considering the timing correlations of neighboring edges decreases the relay times, *i.e.* speeds up the spreading process. This effect is negligible for small weights, but the difference to random relay times reaches



**Figure 7.** The expected normalized average relay times  $\langle \tau_R^* | n \rangle$  (squares),  $\langle \tau_{R1}^* | n \rangle$  (circles), and  $\langle \tau_{R\infty}^* | n \rangle$  (crosses) for the (a) original event sequence, (b) equal-weight link-sequence shuffled and (c) time-shuffled reference models.

$\sim 20$  percent for edges with approximately one call per day. Note the difference between  $\tau_R^*$  and  $\tau_{R1}^*$  for the time-shuffled reference at this point is still negligible, and thus general timing correlations related to the daily pattern cannot be an explanation; rather, this difference is caused by calls triggering further calls. For the equal-weight link-sequence shuffled reference, the relay times  $\tau_R^*$  and  $\tau_{R1}^*$  appear the same. This is because correlations between adjacent links have been destroyed; note that, however, in theory the network-wide daily pattern that is present in this null model could still induce enough correlations to make a difference, but this is not the case. For the time-shuffled reference model, the slowing down effect of the daily pattern disappears completely for the triggered relay times. This happens because for  $\tau_{R1}$ , the triggering infection is more likely to arrive at the edge during daytime when also the focal edge is likely to be more active, and the inter-event times are thus shorter. On the contrary,  $\tau_R$  increases because infections arriving at random points in time may have to wait long times during the nights before crossing the edge.

Finally, we consider the largest scale with the actualized relay times that take network-level spreading process and its timings into account. For normalization, we cannot simply use Eq. (4) as the expected  $\overline{\tau_{R\infty}}$  values for the uniformly random times reference model depend on the topology of the underlying network and other large-scale correlations in the call sequence. Instead, we use the  $\tau_{R\infty}$  values measured from simulations to define the normalized relay time as  $\langle \tau_{R\infty}^* | n \rangle = \langle \tau_{R\infty} | n \rangle / \langle \tau_{R\infty, unif. random} | n \rangle$ . For both the original call sequence and the equal-weight link-sequence shuffled reference, the actualized relay times are shorter than random or triggered relay times, except for the lowest-weight links. The difference grows with increasing weight. Thus, the slowing-down effect of bursts on the actualized relay times is smaller than would be expected just by looking at the links in isolation ( $\tau_R$ ) or additionally the immediate neighborhood of the links ( $\tau_{R1}$ ). One possible explanation for the shorter actualized relay times on strong edges is the fact that the slower an edge is, the more often alternative spreading routes are employed making the progress of spreading faster. Because high-weight edges are much burstier than low-weight edges in relation to the uniform reference model, alternative routes are

used more often for edges with high weights. Even though the normalized actualized relay times are smaller than the corresponding triggered relay times, the qualitative conclusion drawn from the triggered relay times still holds. That is, because the call sequences on edges are bursty, the speed of SI spreading (or equivalently, the duration of fastest temporal paths) is exaggerated if no temporal information about the calls is considered but the calls are assumed to be distributed uniformly random within the observation interval. This is specially true for high-weight edges which slow down more than low-weight edges. Furthermore, the spreading process is not affected much by the daily pattern.

## 5. Discussion

In temporal networks, both the topology of the aggregated network and the timings of interaction events can be crucial in determining how a dynamic process mediated by the network unfolds. We have explored the limiting case of the speed of spreading in the SI model, set up such that an event between an infectious and susceptible individual always transmits the infection. In this process one of the nodes is initially set to the infectious state, and the spread of infection follows the fastest time-respecting paths [9, 7] from this node to all other nodes. Thus the speed of the process sets an upper bound for the speed of any dynamic process that is mediated through the interaction events of the network.

The roles of different types of network correlations in the outcome of a dynamical process can be studied with the help of reference models that switch off selected correlations one by one. For static networks, structural correlations can be removed using the configuration model procedure, while weight-topology correlations can be addressed by shuffling weights. The removal of time-domain correlations in temporal networks is still a fairly new problem, and although several temporal reference models have been designed [9, 5, 1], work in this respect can still be considered to be “in progress”. In addition to the MPC case study, the contribution of this article in more general terms is to formalize and make the use of temporal reference models systematic; we have also analytically calculated some key statistics for the uniformly random times reference model.

The first observation made by using the framework of reference models was that for the MPC network the timings of the sequence of events are as important as the topological structure of the network. Randomizing either the topological structure or call sequences speeds up the spreading process by roughly equal amount. Thus, when estimating the spreading velocity, disregarding the time-domain information and simply aggregating over it results in an error comparable to the error that would result from not taking the network topology into account. In general, this observation points out the importance of taking time-domain information into account when studying dynamics on temporal networks; e.g. a similar slowing down has also been observed for an email network, whereas temporal correlations give rise to faster temporal paths in an air transport network because of optimized scheduling [7].

The second observation made with the reference models was that a large part of the slow dynamics on the MPC network can be attributed to the call sequences of individual links. Thus, the slowing-down effect can be explained by characterizing the timings of the call sequences of links. For this, we have used the concept of relay times [4, 6], and showed how the burstiness and the slow spreading speed compared to the Poisson process are equivalent concepts: the slowing down is due to the bursty

nature of the event sequences. We also found that there is a lot of heterogeneity in the relay times of links even when normalizing by weight; thus not only the broad distribution of link weights but also the distribution of relay times affects the speed of spreading. Further, keeping the exact time sequence of calls on links, rather than considering them uniformly randomly distributed, slows down the high-weight links much more than links with low weights. This points out that the effect of the link weights to the spreading speed would be overestimated without information on event timings.

At the scale of individual links, the spreading speed through a link can be approximated with the random relay time if we only consider information of the events on the focal link. This approximation can be improved by increasing the scale to include neighboring links by adding information about correlation between adjacent event sequences using the triggered relay time [6]. We have constructed a model that shows how such correlations decrease the average relay times; using triggered relay times also corrects for the artificial effect of the circadian and weekly patterns. Finally, we have introduced the actualized relay time, motivated by empirically observed high variations in the likelihood of individual events transmitting the infection during network-scale spreading dynamics. This added level of realism further shortened the relay times. Despite of the overestimation of the relay times when using the two first approximations, the overall picture arising from all the three definitions remains the same; the dominant factor is burstiness.

As the timings of event sequences of individual links have been seen to be the most important factor for spreading dynamics, one might consider mapping the dynamic problem to a static one by defining proper dynamic weights for the links of the aggregated network. This route was recently taken by Miritello *et al.* [6], who defined the dynamic strength of a tie as the probability of it being able to transmit the infection during an SIR spreading process. For the special case of SI dynamics, this probability would always equal one – hence, instead, for the SI process, the interesting quantity is the speed of propagation. One may thus define a dynamical weight that is inversely proportional to the speed *e.g.* as  $n_t = T/\tau_R - 1$  (or, including triggering effects,  $n_t = T/\tau_{R1} - 1$ ) where  $n_t$  is the dynamic weight; for events with uniformly random timings (Eq. 4), this weight would be equal to the number of events on the link with some variation due to fluctuations. However, as we have seen with the actualized transfer times, the real fastest time-ordered paths taken by the spreading process may be considerably different *e.g.* as temporal shortcuts are employed, limiting the accuracy of this approximation.

## 6. Acknowledgments

Financial support from EU’s 7th Framework Program’s FET-Open to ICTeCollective project no. 238597 and by the Academy of Finland, the Finnish Center of Excellence program 2006-2011, project no. 129670, and TEKES (FiDiPro) are gratefully acknowledged. We thank A.-L. Barabási for the data used in this research.

## 7. References

- [1] P. Holme and J. Saramäki. Temporal networks. *arXiv: 1108.1780*, 2011.
- [2] A. Barrat, M. Barthélemy, R. Pastor-Satorras, and A. Vespignani. The architecture of complex

- weighted networks. *Proceedings of the National Academy of Sciences of the United States of America*, 101(11):3747–3752, 2004.
- [3] S. Boccaletti, V. Latora, Y. Moreno, Chavez, and D.-U. Hwang. Complex networks: Structure and dynamics. *Phys. Rep.*, 424:175 – 308, 2006.
  - [4] Alexei Vazquez, Balázs Rácz, András Lukács, and Albert-László Barabási. Impact of non-poissonian activity patterns on spreading processes. *Phys. Rev. Lett.*, 98(15):158702, 2007.
  - [5] M. Karsai, M. Kivelä, R. K. Pan, K. Kaski, J. Kertész, A.-L. Barabási, and J. Saramäki. Small but slow world: How network topology and burstiness slow down spreading. *Phys. Rev. E*, 83(2):025102, 2011.
  - [6] Giovanna Miritello, Esteban Moro, and Rubén Lara. The dynamical strength of social ties in information spreading. *Phys. Rev. E*, 83:045102, 2011.
  - [7] R. K. Pan and J. Saramäki. Path lengths, correlations, and centrality in temporal networks. *Phys. Rev. E*, 84:016105, 2011.
  - [8] Ingve Simonsen, Kasper Astrup Eriksen, Sergei Maslov, and Kim Sneppen. Diffusion on complex networks: a way to probe their large-scale topological structures. *Physica A*, 336:163 – 173, 2004.
  - [9] Petter Holme. Network reachability of real-world contact sequences. *Phys. Rev. E*, 71(4):046119, 2005.
  - [10] Marc Barthélemy, Alain Barrat, Romualdo Pastor-Satorras, and Alessandro Vespignani. Velocity and hierarchical spread of epidemic outbreaks in scale-free networks. *Phys. Rev. Lett.*, 92(17):178701, 2004.
  - [11] J. P. Onnela, J. Saramäki, J. Hyvönen, G. Szabó, D. Lazer, K. Kaski, J. Kertész, and A. L. Barabási. Structure and tie strengths in mobile communication networks. *Proc. Natl. Acad. Sci. U.S.A.*, 104(18):7332–7336, 2007.
  - [12] R.M. Anderson and May R.M. *Infectious Diseases of Humans: Dynamics and Control*. Oxford Science Publications, 1992.
  - [13] H.W. Hethcote. The mathematics of infectious diseases. *SIAM Review*, 42:599, 2000.
  - [14] M.E.J. Newman. Spread of epidemic disease on networks. *Phys. Rev. E*, 66:016128, 2002.
  - [15] Eben Kenah and James M. Robins. Second look at the spread of epidemics on networks. *Phys. Rev. E*, 76:036113, 2007.
  - [16] Martina Morris and Mirjam Kretzschmar. Concurrent partnerships and the spread of hiv. *AIDS*, 11(5):641648, 1997.
  - [17] Luis E. C. Rocha, Fredrik Liljeros, and Petter Holme. Information dynamics shape the sexual networks of Internet-mediated prostitution. *Proc. Natl. Acad. Sci. U.S.A.*, 107(13):5706–5711, 2010.
  - [18] Luis E. C. Rocha, Fredrik Liljeros, and Petter Holme. Simulated epidemics in an empirical spatiotemporal network of 50,185 sexual contacts. *PLoS Comput Biol*, 7(3):e1001109, 2011.
  - [19] A. L. Barabási. The origin of bursts and heavy tails in human dynamics. *Nature*, 435:207–211, 2005.
  - [20] José Luis Iribarren and Esteban Moro. Affinity paths and information diffusion in social networks. *Social Networks*, 33:134–142, 2011.
  - [21] Michael Molloy and Bruce Reed. A critical point for random graphs with a given degree sequence. *Random Structures & Algorithms*, 6:161180, 1995.
  - [22] H.-H. Jo, M. Karsai, J. Kertész, and K. Kaski. Circadian pattern and burstiness in mobile phone communication. *arXiv:1101.0377v2*, 2011.
  - [23] K.-I. Goh and A.-L. Barabási. Burstiness and memory in complex systems. *Europhys. Lett.*, 81:48002, 2008.
  - [24] Ye Wu, Changsong Zhou, Jinghua Xiao, Jürgen Kurths, and Hans J. Schellnhuber. Evidence for a bimodal distribution in human communication. *Proc. Natl. Acad. Sci. U.S.A.*, 107(44):18803–18808, 2010.
  - [25] Byungjoon Min, K.-I. Goh, and Alexei Vazquez. Spreading dynamics following bursty human activity patterns. *Phys. Rev. E*, 83(3):036102, 2011.
  - [26] Julián Candia, Marta C González, Pu Wang, Timothy Schoenharl, Greg Madey, and Albert-László Barabási. Uncovering individual and collective human dynamics from mobile phone records. *J. Phys. A: Math. Theor.*, 41(22):224015, 2008.
  - [27] F.W. Steutel. Random division of an interval. *Statistical Neerlandica* 21, 21:231–244, 1967.
  - [28] Herbert A. David and H.N. Nagaraja. *Order Statistics*. Wiley-Blackwell, third edition, 2003.
  - [29] Z. Eisler, I. Bartos, and J. Kertész. Fluctuation scaling in complex systems: Taylor’s law and beyond. *Advances in Physics*, 57(1):89–142, 2008.
  - [30] Gautier Krings, Francesco Calabrese, Carlo Ratti, and Vincent D Blondel. Urban gravity: a model for inter-city telecommunication flows. *J. Stat. Mech.*, 2009(07):L07003, 2009.

- [31] R. K. Pan, M. Kivelä, J. Saramäki, K. Kaski, and J. Kertész. Using explosive percolation in analysis of real-world networks. *Phys. Rev. E*, 83:046112, 2011.
- [32] P.A.P. Moran. The random division of an interval. *Supplement to the Journal of the Royal Statistical Society*, 9:92–98, 1947.

## Appendix A. Reference models and static and dynamic properties

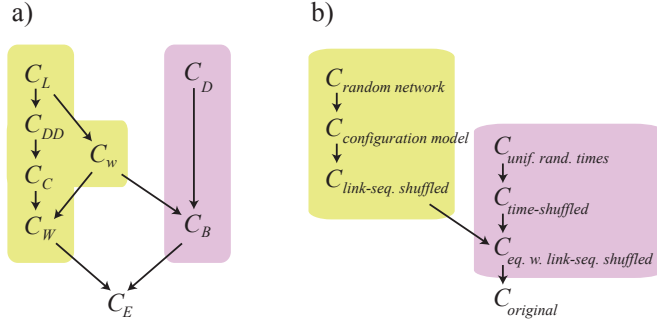
Empirical systems are usually characterized by properties which are extremely unlikely to be found in random systems. One way of studying static empirical networks is to compare them to an ensemble of graphs which possess some subset of the properties of the empirical system but which are otherwise maximally random. These ensembles are often called as *reference* or *null models*, and they can be used to perform “controlled experiments” to see which properties of the data are relevant for some observed phenomenon. The most convenient way of generating samples from random ensembles is to use shuffling methods which retain the selected set of features of the original system but randomize and destroy all other unrelated properties. Then, one can observe how the quantities of interest differ in the reference ensemble as compared to the empirical system. In this section, we define a reference model framework for temporal networks by formally defining the properties that the reference models modify.

A temporal network can be described by a set of nodes  $v \in V$  which are connected by a sequence of events,  $\mathcal{E} = \{e_1, \dots, e_N\}$ , occurring between the nodes during the observation period  $[0, T]$ . Each event can be defined as a quadruplet of the form  $e \equiv (u, v, t, d)$ , where  $u$  and  $v$  are the initiator and receiver of the event,  $t \in [0, T]$  denotes the execution time and  $d$  is the duration. In this study, we analyze mobile phone calls which have durations, but for simplicity we consider them instantaneous as their duration is irrelevant for the investigated processes. Consequently, the definition of an event is simplified to a triplet form  $e \equiv (u, v, t)$ . A sequence of such events can be used to construct an aggregated network  $G(\mathcal{E})$  where a link appears between nodes who participate in a common event and link weights can be defined as the total number of events  $n$  taking place between connected nodes during the examined time period.

Formally, a property of an event sequence can be any function  $c$  that takes an event sequence as an argument and returns the property. Using this function the set of all possible event sequences can be divided into two subsets, one with sequences having the property,  $C(\mathcal{E}) = \{\mathcal{E}' | c(\mathcal{E}) = c(\mathcal{E}')\}$  and the rest of the sequences that do not have the property. The reference model related to a property is an ensemble of event sequences containing all the sequences in  $C(\mathcal{E})$  where each sequence is equally probable. At the crudest level, the properties of an event sequence can be divided into static properties which can be detected from the aggregated network (*i.e.* we can write  $c(\mathcal{E}) = c'(G(\mathcal{E}))$ ) and to temporal properties for which we also need to consider the time stamps of the events.

Let us first concentrate on the static properties. Without any constraints all of our event sequences have the same set of nodes  $V$  and the ensembles of possible graphs induced by these sequences are limited to graphs with the same node set. The number of events in each sequence is  $N$  and thus the number of edges in the graphs vary from 1 to  $N$ . The first constraint we want to consider is the *number of links*  $c_L(\mathcal{E}) = L(G(\mathcal{E}))$ , where  $L$  is a function returning the number of links in a graph. Now the ensemble of graphs induced by the ensemble of event sequences correspond

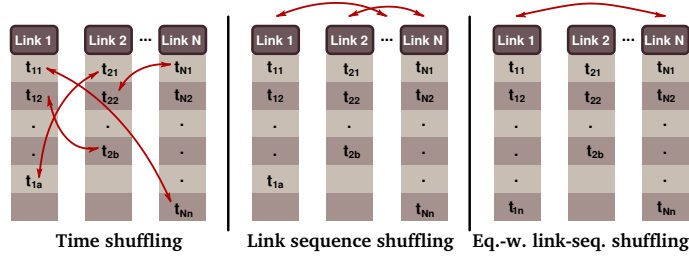
to an ensemble of Erdős-Rényi (E-R) random graphs with given numbers of nodes and edges. E-R graphs have a binomial degree distribution, instead of the fat-tailed distributions commonly observed in data. Consequently, the next step is to limit the graphs to have exactly the same *degree sequence* as the empirical graph. That is, we define a property  $c_{DD}(\mathcal{E}) = k(G(\mathcal{E}))$ , where  $k$  is the function returning the degree sequence of a graph. This ensemble is called the configuration model. The full topology of the network is returned if in addition to the degree sequence we include the topological configurations and restore the original connection structure. That is, we keep the property of *topological configurations* with the constraint  $c_C(\mathcal{E}) = A(G(\mathcal{E}))$ , where  $A$  is the unweighted adjacency matrix of the graph.



**Figure A1.** Illustration of the hierarchical structure of (a) the properties discussed in the text, and (b) the randomized reference models. Arrows denote the inclusions of different ensembles in a way that there is an arrow from  $A$  to  $B$  if  $A \supseteq B$ . Static properties and reference models are located on the left-hand side, while temporal properties and reference models are on the right.

Until now, we have neglected the edge weights of the network, *i.e.* the number of events taking place between node pairs, and in our ensembles the weight distribution has been maximally random, or binomial. If we define link weights as the numbers of events on links, we can constrain the *weight distribution*  $c_w(\mathcal{E}) = N_w(G(\mathcal{E}))$ , where  $N_w(G)(n) = \sum_{i,j} \delta(G_{i,j}, n)$  is the count of weights in the network represented by the weight matrix  $G$  and  $n$  denotes the number of events on a link. Now the number of edges having a given weight is determined, but the weights are distributed randomly between edges. Finally, to restore all static properties, we can include information about which weight corresponds to which edge as we define the property corresponding to the *weight-topology correlations* as the weight matrix of the aggregated graph,  $c_W(\mathcal{E}) = G(\mathcal{E})$ . As we can conclude from the way we added increasingly strict constraints for the ensemble, the reference models present a hierarchical structure and the hierarchy relations can be expressed with the sets of event sequences having non-zero probability in the ensembles:  $C_L \supseteq C_{DD} \supseteq C_C \supseteq C_W$  and  $C_L \supseteq C_w \supseteq C_W$ . This structure is also illustrated on Fig A1 a).

The study of temporal inhomogeneities in networks has gained popularity only recently, and the related reference models are not well established. Next, we define three temporal properties which are used in this paper. Without any temporal constraints, the times of the events in the ensembles are scattered uniformly at random on the interval  $[0, T]$ . First, we want to limit the ensemble in such a way that at least the network-wide frequency patterns in the event times are preserved. To do



**Figure A2.** Call sequence shuffling procedures corresponding to null models which only remove temporal inhomogeneities.

so, we retain the *global time stamp sequence*  $c_D(\mathcal{E}) = \mathcal{T}$ , where  $\mathcal{T} = \{t_1, \dots, t_N\}$  is the sequence of the time stamps of the original event sequence  $\mathcal{E}$ . In a mobile phone network, the frequency patterns of the time stamps  $\mathcal{T}$  reflect the daily and weekly fluctuations of the calling frequency, as on average people tend to place more calls at day time and less calls at night with a slight variation due to different weekdays. Going one step further, we would like to keep the distribution of *the times of event sequences on links*  $c_B(\mathcal{E}) = N_S(\mathcal{E})$  with  $N_S(\mathcal{E})(s) = \sum_{i,j} \delta(S_{i,j}, s)$ , where  $S_{i,j} = \{t_{k_1}, \dots, t_{k_n} | u_{k_1} = \dots = u_{k_n} = i, v_{k_1} = \dots = v_{k_n} = j\}$  is the sequence of times of events taking place between the nodes  $i$  and  $j$ . The link sequences in the MPC network are known to be bursty. Restoring the full link sequences retains all the characteristics of the times of event sequences on links, including inter-event times and burstiness. The event sequences on links can be perceived as “temporal link weights”, *i.e.* sets of time stamps associated with each link instead of the simple scalar value of static weights. Similarly to static link weights, if in addition to retaining their overall distribution we want to keep the exact locations of each link sequence in the network, then the complete original event sequence  $\mathcal{E}$  is returned. Thus, for completeness we define a property of retaining the *link sequence-topology correlations* by  $c_E(\mathcal{E}) = \mathcal{E}$ . The inhomogeneities that are captured by  $c_E$  (and not by any other property defined here) include burst-topology correlations and between-link correlations such as events causing other events in neighboring links, or for example some interlinked sets of nodes being more active than others at night. For the temporal reference models, the hierarchy relations are  $C_D \supseteq C_B \supseteq C_E$ , and between temporal and static references they are  $C_w \supseteq C_B, C_W \supseteq C_E$ .

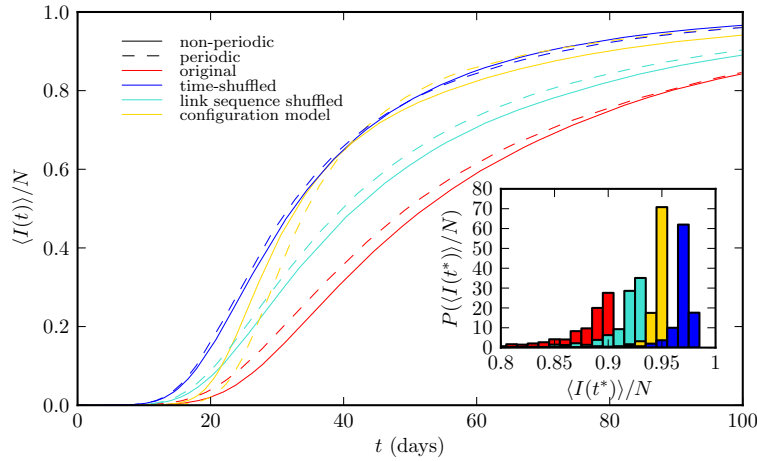
Any two properties  $c_A$  and  $c_B$  can be combined by defining that  $c_{A+B}(\mathcal{E}) = \{c_A(\mathcal{E}), c_B(\mathcal{E})\}$  or equivalently by having  $C_{A+B} = C_A \cap C_B$ . The different properties can be combined in any way to create reference models, but some combinations are redundant because the reference models can contain each other. All the reference models listed in the main text can now be defined as such combinations (see Sec. 2.2 and Table 1.)

## Appendix B. Role of boundary and initial conditions

To ensure that our results are not heavily affected by the periodic temporal boundary conditions that we applied, we have repeated some of our reference model calculations without periodic boundaries, using only the available finite time window of 120 days. Here, the infection was initiated from random node at a random point in time within



the first 20 days, and the process was run for 100 days. Repeating the measurements for  $10^3$  times with different initial conditions, the average infection curves (solid lines in Fig.B1) qualitatively show the same behavior as for the periodic boundary conditions (dashed lines in Fig.B1). The overall effects are minimal, but the periodic boundary conditions slightly fasten the spreading process for the reference models having the original link-sequence distribution. For the time-shuffled event sequence where bursty temporal correlations do not play a role, the two curves (with and without periodic boundaries) overlap almost perfectly. Since without periodic boundary conditions the system does not reach full prevalence during the examined time period, we recorded the distribution of the infection curve values  $\langle I(t = 100) \rangle / N$  at the end of the time window (see Fig.B1 inset) in order to see the effects of random initial conditions. Although there is variance and the tails of the distributions overlap, their order is the clearly same as for the means of the curves.



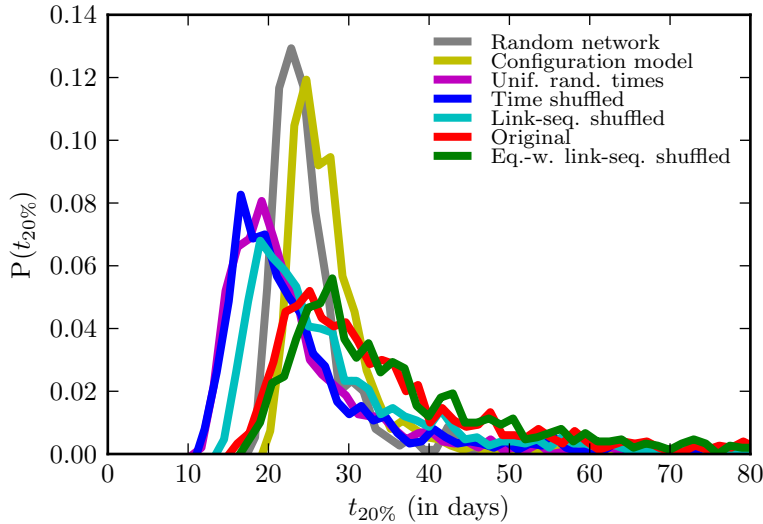
**Figure B1.** Comparison between spreading dynamics with (dashed line) and without (solid line) periodic temporal boundary conditions. Distribution of  $\langle I(t^*) \rangle / N$  at time  $t^* = 100$  days of different null models without periodic boundary conditions is shown in the inset.

The initial conditions, *i.e.* the first infected node and the starting time, can have a large impact on the spreading process. The topological position and the activity of the initially infected node strongly influences the speed of a realization of the process, as the infection lifts off faster when it is initiated from a well-connected hub with a high level of activity. Such differences are reflected in the variation of the distributions of the times to reach prevalence levels of 20% and 100%, as shown in Fig.B2 and in the inset of Fig.1, respectively. The effects of the initial conditions can be quantified by the width of the distributions, as summarized in Table B1, where the average and the error of the mean values of the distributions of times to reach prevalences of 20% and 100% are shown for all models. Especially for early-stage distributions in Fig.B2, the reference models that retain more inhomogeneities have a larger spread.

A more sophisticated picture can be drawn by recording the density of the infection curves in a time-prevalence space. For computing the density, this space can be divided into bins of equal size; then, the number of times when an infection curve passed through it can be counted for each bin cell. In Fig.B3 we show the normalized density figures of three selected reference models together with the average curves. In

EVENT SEQUENCE	$\langle t_{20\%} \rangle$	$\langle t_{100\%} \rangle$
Original	$38.242 \pm 0.741$	$677.670 \pm 1.008$
Eq.-w. link-seq. shuffled	$41.206 \pm 0.831$	$597.267 \pm 1.592$
Time shuffled	$26.486 \pm 0.527$	$402.400 \pm 1.141$
Unif. rand. times	$24.962 \pm 0.390$	$401.649 \pm 1.113$
Link-seq. shuffled	$30.676 \pm 0.552$	$495.981 \pm 1.002$
Configuration model	$29.359 \pm 0.327$	$446.621 \pm 0.949$
Random network	$26.643 \pm 0.291$	$411.803 \pm 1.139$

**Table B1.** Average time for reaching prevalence levels of 20% and 100% infection, and error of the mean values for each model.

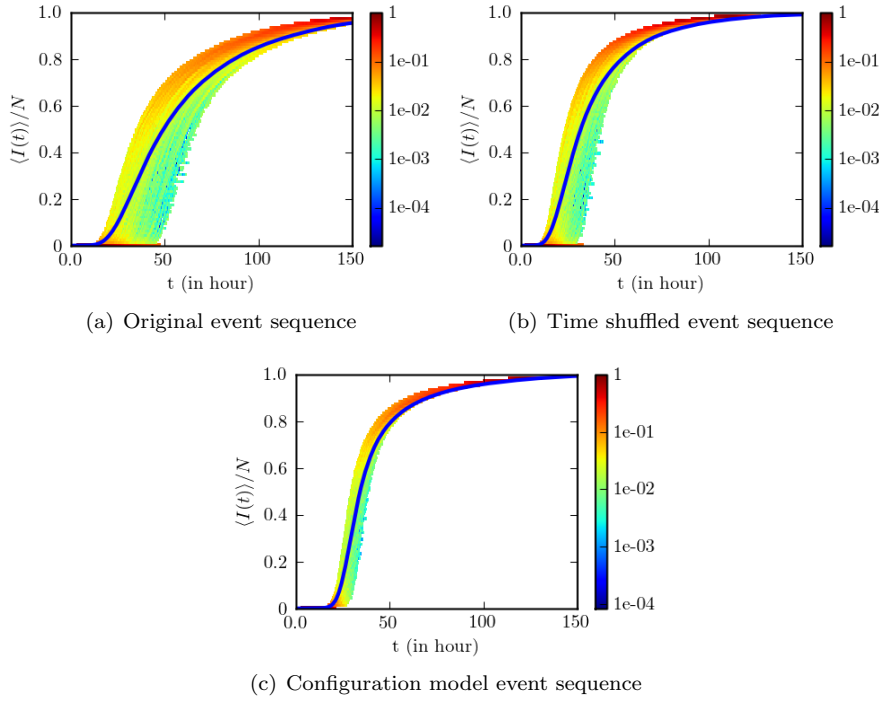


**Figure B2.** Distributions for times to reach 20% prevalence.

order to display the regime where the vast majority of the infection curves are running, we have used quantiles for the distributions for each row and show only the 75% of the curves symmetrical around the median. This demonstrates that even though initial conditions cause an impact on the spreading evolution, most of the curves are not too far from the ensemble average which gives then a fairly good estimate for the actual behaviour. The effect of heterogeneities increasing the spread is also visible here, as the dispersion of the individual curves is decreasing around the average and the 75% quantile area becomes more narrow for reference models with less constraints.

### Appendix C. Network size dependence

In order to study the dependence of the speed of SI spreading process on the system size, we need to extract subnetworks of the MPC network. Here, we take use of the fact that MPC networks display geographical embedding [30], and divide the users into subsets by considering in which city they live according to the postal code of their subscriptions [31]. We then determine the largest connected component (LCC)

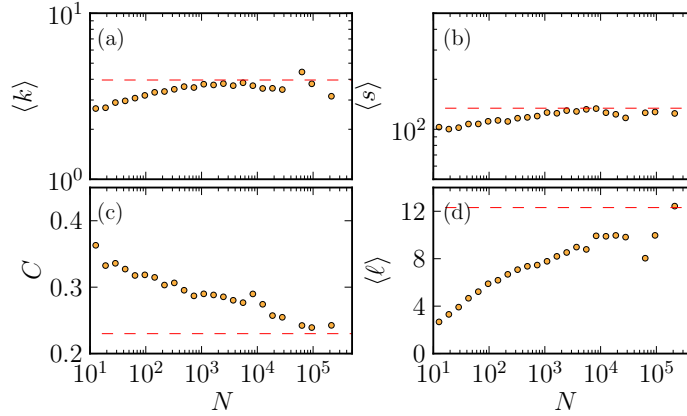


**Figure B3.** Density and average of infection curves in the time-infection space for different null models. (a) Original event sequence, (b) time shuffled event sequence and (c) configuration model. On each figure the 75% quantiled distributions are shown together with the average infection curves (solid line).

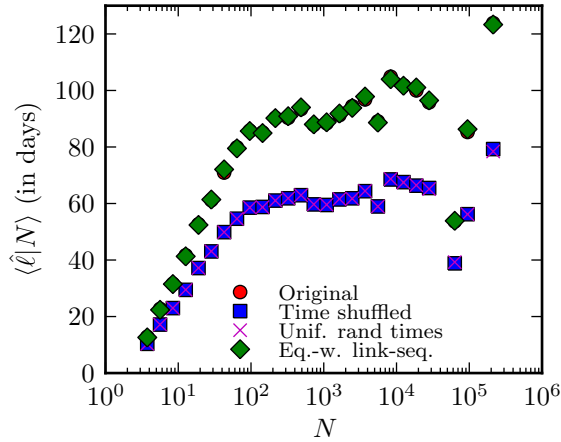
of each city network and construct the event sequence data of that city by selecting all the calls between the users in the LCC. The above procedure yields 7569 subnetworks, whose sizes vary over five orders of magnitude. Note that the information about the postal code of subscription is available for only about half of the users, and hence the amount of users in all subnetworks is less than in the entire MPC network.

MPC networks inside cities form natural subsystems of the whole MPC network [31, 30]. There is no reason to believe that the social systems of different sizes are similar; instead, it might be possible that the interaction patterns of people in the entire country might have very different characteristics as compared to the interaction patterns inside a small city. Further, the size of the city might also have some effect on the patterns and correlations of interaction between users. We start by inspecting the properties of the corresponding static networks.

The average degree  $\langle k \rangle$  of the subnetworks initially increases with  $N$ , but for larger systems it saturates when the value for the entire MPC network is reached [Fig. C1 (a)]. The average strength  $\langle s \rangle$ , i.e. the average number of calls in which nodes participate, also behaves similarly [Fig. C1 (b)]. These properties suggest that although there is a large variation in the sizes of the subnetworks, their fundamental properties are similar. Typical to social networks, the clustering coefficients  $C$  of the entire MPC network and the subnetworks are much higher than would be expected from Erdős-Rényi random networks with corresponding size and average degree [Fig. C1 (c)]. The



**Figure C1.** Variation of the static properties of the subnetworks with the system size,  $N$ . Plot of the (a) average degree  $\langle k \rangle$ , (b) average distance  $\langle \ell \rangle$ , (c) average strength  $\langle s \rangle$ , and the (d) clustering coefficient  $C$ , of the subnetworks, as a function of their size. The dashed line represents the corresponding value for the entire MPC network.



**Figure C2.** The network-wide average temporal distance between two nodes (time to reach  $i$  from  $j$ ) as a function of the number of nodes  $N$  in the network for the original data and several null models. The original curve overlaps with the equal-weight link-sequence shuffled model; the time-shuffled and uniform random times curves also practically overlap. All curves are averaged over  $10^3$  initial conditions.

value of  $C$  decreases with increasing network size, reaching the value of the entire MPC network for large  $N$ . Fig. C1 (d), shows the variation of the average shortest path length in the subnetworks  $\ell$  with  $N$ . As expected, we find that the *static* MPC network is a small-world system where (roughly)  $\langle \ell \rangle \propto \ln N$ . Note that when  $N > 2 \times 10^3$ , we calculate  $\langle \ell \rangle$  by sampling the distance from  $10^3$  randomly chosen nodes to all other nodes in the network.

Next, we perform the SI spreading process for each of the subsystems, using

both the original event sequence and the temporal reference models. We simulate the process until all the nodes in the system are infected and observe the infection time of each node. We then measure the average time at which nodes get infected in a city as a function of the city size  $N$  [Fig C2]. This average infection time is same as the average temporal distance (duration of the fastest paths) of the system. The temporal distances first show a fast increase, followed by slower increase. Importantly, for all system sizes, the temporal distances for the original data and the null models show a similar relationship. This indicates that the observed results are unlikely to be an artifact associated with the specific system, that is, the full MPC network.

#### Appendix D. $\tau_R$ for the uniformly random times reference model with periodic boundary conditions

In this Appendix we are going to calculate the inter-event time and random relay time distributions for the uniformly random times reference model. Further, we are going to derive formulas for the expected average relay time and variance for the average relay time.

In the uniformly random times reference model the sequences of events of edges do not depend on each other and we can calculate all statistics just by considering a single edge at a time. That is, we have an edge with  $n$  events, each taking place uniformly at random in the interval  $[0, T]$ , and denote the time of the  $i$ th event (ordered by time) by  $t_i$ .

The time of  $i$ -th event is distributed like the  $i$ -th order statistics of the uniform probability distribution  $t_i/T \sim U_{(i)}$ . That is,  $t_i/T$  follows the beta distribution [28]

$$p(t_i) = \frac{(t_i/T)^{i-2}(1 - t_i/T)^{n-i+1}}{B(i-1, n-i)}, \quad (\text{D.1})$$

where  $B(i-1, n-i) = \int_0^1 x^{i-2}(1-x)^{n-i+1}dx$  is the beta function.

In order to be able to calculate  $\tau_R$ , we need the distribution for the inter-event times  $\tau_i$ , rather than the times of the events. With periodic temporal boundary conditions in place, we have

$$\tau_i = \begin{cases} T - t_n + t_1 & \text{if } i = 1 \\ t_i - t_{i-1} & \text{if } i \neq 1 \end{cases}. \quad (\text{D.2})$$

Note that the values of  $t_i$  are not independent of each other. However, from Ref. [28], we know that if we define a new random variable  $W_{rs} = U_{(s)} - U_{(r)}$  we get

$$p_{W_{rs}}(w) = \frac{w^{s-r-1}(1-w)^{n-s+r}}{B(s-r, n-s+r+1)} \quad (\text{D.3})$$

Applying Eq. D.3 to Eq. D.2, the inter-event time  $\tau_i$  distribution takes the form

$$p(\tau_i) = \begin{cases} \frac{(\tau_i/T)^{n-2}(1-\tau_i/T)}{B(n-1, 2)} & \text{if } i = 1 \\ \frac{(1-\tau_i/T)^{n-1}}{B(1, n)} & \text{if } i \neq 1 \end{cases}. \quad (\text{D.4})$$

The expected value of  $\tau_i$  thus becomes

$$\langle \tau_i \rangle = \begin{cases} \frac{2T}{n+1} & \text{if } i = 1 \\ \frac{T}{n+1} & \text{if } i \neq 1 \end{cases}, \quad (\text{D.5})$$

and the variance is

$$\text{Var}(\tau_i) = \begin{cases} \frac{2(n-1)}{(n+2)(n+1)^2} T^2 & \text{if } i = 1 \\ \frac{n}{(n+2)(n+1)^2} T^2 & \text{if } i \neq 1 \end{cases}. \quad (\text{D.6})$$

Note that  $\tau_1$  is longer than all the other inter-event times. At first this might seem to be a bias in the reference model. However, if one would change the labeling scheme of the event times such that  $t_1$  would be the first event after some point  $x \in [0, 1]$  and use the periodic boundary condition in labeling all other events such that  $t_{n-1}$  would be the last event before  $x$ , this would not change the distribution of  $\tau_i$ . This might seem paradoxical, but it just reflects the fact that if we select any point  $x$  in the time interval, we are more likely to hit long inter-event times than short inter-event times. That is, selecting the point  $x = 0$  as basis of our “coordinate system” causes the first inter-event time to be longer than the others.

Equipped with the functional form of the probability density function of the distribution of  $\tau_i$ , we can calculate the expected value of  $\tau_R$  by using the Eq. 3:

$$\langle \overline{\tau_R} \rangle = \left\langle \frac{\sum_{i=1}^n \tau_i^2}{2T} \right\rangle = \frac{\sum_{i=1}^n \langle \tau_i^2 \rangle}{2T} = \frac{\langle \tau_1^2 \rangle + (n-1) \langle \tau_2^2 \rangle}{2T}. \quad (\text{D.7})$$

Now noting that  $\langle \tau_i^2 \rangle = \langle \tau_i \rangle^2 + \text{Var}(\tau_i)$  and plugging in Eq D.5 and Eq. D.6 we get

$$\langle \overline{\tau_R} \rangle = \frac{T}{n+1}. \quad (\text{D.8})$$

The calculations can be made easier if we only want to calculate the average relay times and do not care about the distribution of  $\tau_i$ . We can choose a random event  $i$  and set our time reference in a way that  $t_i = 0$  and still apply the periodic boundary conditions. The set of inter-event times is retained in this transformation, and only the order of the times is lost. Further, the other  $n-1$  events (not including event  $i$ ) are uniformly randomly distributed over the time interval  $[0, T]$ . Thus, the problem is of finding the distribution of the average relay time can be solved by solving the problem of dividing the unit interval into  $n$  intervals by  $n-1$  uniformly random points. In Ref. [32] the  $l$ 'th moment of the sum of the squares of the interval sizes was given as

$$\langle (\sum_{i=1}^n (\tau_i/T)^2)^l \rangle = \frac{(n-1)!!}{(n+2l-1)!} \sum_{j_1+\dots+j_n=l} \frac{(2j_1)! \dots (2j_n)!}{j_1! \dots j_n!}. \quad (\text{D.9})$$

Now the  $l$ 'th moment of the average relay time can be calculated by using Eq. D.9 with

$$\langle \overline{\tau_R}^l \rangle = \left(\frac{T}{2}\right)^l \langle (\sum_{i=1}^n (\tau_i/T)^2)^l \rangle. \quad (\text{D.10})$$

Using Eqs. D.9 and D.10 we obtain the correct average relay time of Eq. D.8, and values for the second moment

$$\langle \overline{\tau_R}^2 \rangle = \frac{n+5}{(n+3)(n+2)(n+1)} T^2, \quad (\text{D.11})$$

and the variance

$$\text{Var}(\overline{\tau_R}) = \frac{n-1}{(n+1)^2(n+2)(n+3)} T^2. \quad (\text{D.12})$$